Sufficient conditions for two-dimensional localization by arbitrarily weak defects in periodic potentials with band gaps

Arthur Parzygnat
Department of Physics, Queens College, Flushing, New York 11367, USA

Karen K. Y. Lee
Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Yehuda Avniel
Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

Steven G. Johnson*
Department of Mathematics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA
(Received 23 February 2010; revised manuscript received 25 March 2010; published 28 April 2010)

We prove, via an elementary variational method, one-dimensional (1D) and two-dimensional (2D) localization within the band gaps of a periodic Schrödinger operator for any mostly negative or mostly positive defect potential, V, whose depth is not too great compared to the size of the gap. In a similar way, we also prove sufficient conditions for 1D and 2D localization below the ground state of such an operator. Furthermore, we extend our results to 1D and 2D localization in d dimensions; for example, by a linear or planar defect in a three-dimensional crystal. For the case of D-fold degenerate band edges, we also give sufficient conditions for localization of up to D states.

DOI: 10.1103/PhysRevB.81.155324 PACS number(s): 73.20.Hb

I. INTRODUCTION

Localization by impurities in periodic potentials with spectral gaps (band gaps) is a central topic in solid-state physics and semiconductor devices[1] and it has direct analogs for other propagating-wave systems, such as for photonic crystals in electromagnetism.[2] We prove that, in one and two dimensions (1D and 2D), localized solutions must arise in the gaps of a periodic Schrödinger operator for any "mostly" negative or mostly positive defect potential whose depth is not too great compared to the size of the gap. To our knowledge, this is the first rigorous sufficient condition of this sort in 2D, aside from informal arguments based on effective-mass free-electron models close to quadratic gap edges,[3] extending an earlier theorem for 1D localization in gaps,[4] and is quite different from the many rigorous asymptotic gap-localization results in the limit of very *strong* defect potentials.[5–10] In addition to localization in gaps, we also prove 1D and 2D localization below the ground state of a periodic potential for any mostly negative defect potential, extending earlier known results for localization in vacuum for any mostly negative potential,[11–13] localization in the periodic case but for a strictly nonpositive defect potential,[14,15] and 1D localization for mostly negative defects.[4] Furthermore, we extend our results to 1D and 2D localization in d dimensions; for example, establishing localization for a linear or planar defect in a three-dimensional (3D) crystal. For the case of D-fold degenerate band edges, we show localization of D bound states for definite-sign defect potentials and more generally relate the number of localized modes to the signs of the eigenvalues of a $D \times D$ matrix. Our proofs rely only on elementary variational eigenvalue bounds, generalizing an approach developed in Ref. 13.

One-dimensional localization in vacuum by an arbitrary attractive potential is easy to prove by a variational method, at the level of an undergraduate homework problem,[16] while a minimum nonzero depth for a potential well is required for localization in vacuum in three dimensions.[17] The 2D case is more challenging to analyze. The first proof of 2D localization in vacuum by an arbitrary attractive potential was presented in 1976 by Simon[11] using more sophisticated techniques that also lead to asymptotic properties of the bound states. An elementary variational proof in vacuum was proposed by Picq[18,19] (and was adapted to Maxwell's equations in optical fibers by Bamberger and Bonnet[19]), while a different variational proof was independently developed by Yang and de Llano.[13] An informal asymptotic argument utilizing properties of the vacuum Green's function was presented by Economou.[12] On the other hand, in the case of periodic potentials, rigorous results for the existence of bound states from weak defect potentials V are more limited. Frank *et al.*[15] analyzed the case of potentials $V \leq 0$ localizing at energies below the ground state of a 2D periodic Schrödinger operator; they not only proved that a localized state exists, but also bounded its energy in terms of a related vacuum localization problem. Here, we use a different technique, based on the variational method of Ref. 13, to prove existence of localized modes in 1D and 2D for any indefinite-sign but "mostly" negative defect potential V. This is closely related to our generalization of Ref. 13 to index guiding in the periodic Maxwell's equations.[20] (Indefinite-sign localization was also considered for *discrete* Schrödinger operators.[21]) As for localization in arbitrary gaps, however, we are not aware of published rigorous results in 2D that are valid for weak V. Prodan[4] proved that arbitrarily weak de-

fects localize states in 1D gaps using an asymptotic Birman-Schwinger technique similar to Ref. 11, imposing a mostly negative or positive condition on the defect potential identical to the one we use below; Prodan's result also applies to localization below the ground state of a periodic 1D potential. Various authors have shown that for *strong* defect potentials, those of the form λV where $\lambda \to \infty$, there exists a bound on the number of eigenvalues crossing the gap.[5–9] Localization has also been proved in the limit of high-order gaps in 1D.[22] Another common, albeit somewhat informal, approach to gap localization is to consider localization for energies close to a nondegenerate quadratic band edge, making an effective-mass approximation and then quoting the results for vacuum.[3] Our proof of localization in gaps is nonasymptotic, does not assume a particular form of the band edge, and is an extension of the elementary variational technique of Ref. 13 for localization in vacuum. The trick behind the variational proof is to take the midgap energy, E_g, of the perturbed Schrödinger operator, H, and transform the question of energies near E_g into an extremal eigenvalue problem. There are two typical ways to do this. One is to consider $(H-E_g)^{-1}$, which seems closely related to the Green's function method of Ref. 12 and to the Birmann-Schwinger condition,[23] but such an operator is hard to evaluate explicitly and mainly lends itself to asymptotic analyses. Another method is to consider $(H-E_g)^2$, recently used for another variational-type localization proof by Ref. 24, and it is this method that we adopt here. The same techniques are used in numerical methods to iteratively compute eigenvalues in the interior of the spectrum of a large matrix, where $(H-E_g)^{-1}$ corresponds to well-known shift-and-invert methods[25] and where $(H-E_g)^2$ has also been used (but is computationally suboptimal because it squares the condition number).[26,27] Other possible techniques[9,22,28] have been suggested to us[29] for proving such a theorem, but we are not aware of any published results for this problem other than the 1D result of Ref. 4. Localization by weak defects is also related to self-focusing of solitons by nonlinearities, which was recently considered for spectral gaps in periodic potentials.[30]

The contents of the individual sections are summarized as follows. Section II provides a simple variational proof of the fact that any arbitrary mostly negative and sufficiently localized defect induces at least one bound state below the infimum of the spectrum of a 2D periodic Schrödinger operator. The 1D case is proved in a similar fashion. Section III gives a generalization of that result by allowing the unperturbed Hamiltonian to be periodic in d dimensions and our defect potential to localize in two (or one) dimensions, but be periodic (with the same periodicity as the unperturbed potential) in the other $d-2$ (or $d-1$) dimensions. Section IV presents our main results: it gives a sufficient condition for the existence of a bound state in a spectral gap of a 2D periodic Schrödinger operator. The results from Sec. III can be applied to this case as well to allow periodicity in any arbitrary extra number of dimensions (although the solution is only localized in two of these dimensions). The case of bound states in a gap confined along one dimension is proved in a similar way. Section V describes the case of degenerate band edges as well as some other possible generalizations. Section VI closes with some concluding remarks and possible future directions.

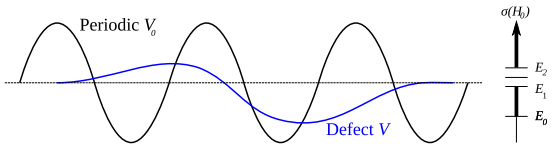

FIG. 1. (Color online) Example of a periodic potential V_0 and defect potential V with the spectrum $\sigma(H_0)$ located at the right. The spectrum of the unperturbed Schrödinger operator has an infimum E_0 and may also have a gap in the set (E_1, E_2).

II. BOUND STATES BELOW THE GROUND STATE OF A PERIODIC POTENTIAL

Notation. In this section, unless otherwise stated, the symbol $\int f$ will stand for the integral over all of $\mathbb{R}^2$ of the real-valued function f.

The proof of the following will be a little simpler than the one in Sec. IV. However, the ideas used here are almost the same as those of the proof in Sec. IV, so it is hoped that after going through this proof, the reader will easily follow the latter. Note that the theorem in this section has also been proved by other methods, but only for defect potentials that are strictly nonpositive.[14,15] This theorem will be slightly generalized to allow for a defect potential that is localized in two dimensions, but has an arbitrary periodicity in all other dimensions in Sec. III.

A. Problem statement

Suppose we start with an unperturbed Hamiltonian

$$H_0 = -\nabla^2 + V_0, \tag{1}$$

where V_0 is a periodic potential (possible generalization to nonperiodic potentials is discussed in Sec. V A), which has a minimum-energy eigenvalue E_0 with at least one (degeneracy will be explored in Sec. V B) corresponding "generalized" eigenfunction ψ_0 of the Bloch form:[1] a periodic function with the same periodicity as V_0 multiplied by $e^{i\mathbf{k}\cdot\mathbf{x}}$, giving a bounded $|\psi_0|$. We introduce a localized, indefinite (varying sign) defect, V, giving a new Hamiltonian $H=H_0+V$, satisfying the two conditions

$$\int V|\psi_0|^2 < 0 \quad \text{and} \tag{2}$$

$$\lim_{r\to\infty} V(r,\theta) = 0. \tag{3}$$

Such a V_0 and V are shown schematically in Fig. 1 along with a typical spectrum showing E_0 (and also possibly a gap, which is considered in Sec. IV). We prove that conditions (2) and (3) are sufficient to guarantee the existence of a bound state with energy lower than E_0. [Strictly speaking, in order for an energy $<E_0$ to be a "bound" state, the potential V must be sufficiently localized so as to not alter the essential spectrum (energies of nonlocalized modes) of H_0; Eq. (3) does not appear to be strong enough. Weyl's theorem provides a sufficient condition of relatively compact V, such as the Kato class of mostly square-integrable potentials, which seems to include most realistic localized defects.[31,32]] The proof is essentially a generalization of that in Ref. 13.

B. Proof of 2D localization

Recall that the variational principle (or min-max principle) states that the Rayleigh quotient $\langle\psi, H\psi\rangle/\langle\psi,\psi\rangle$, where $\langle\psi,\phi\rangle \equiv \int \psi^* \phi$ denotes the inner product between ψ and ϕ (and $\|\psi\|^2$ denotes $\langle\psi,\psi\rangle$), for any trial function ψ (in the appropriate Sobolev space) is an upper bound for the ground-state eigenvalue of a Hermitian operator H.[33–35] Therefore, to prove the existence of an eigenvalue less than E_0 (and thus a bound state), it suffices to find a ψ such that

$$E[\psi] \equiv \frac{\langle\psi,(H-E_0)\psi\rangle}{\langle\psi,\psi\rangle} < 0. \tag{4}$$

The key point is to find a trial function that will work even for an arbitrarily small defect V. Motivated by Ref. 13, we use the following trial function parametrized by a positive number α:

$$\psi = \psi_0 \gamma, \quad \text{where} \quad \gamma = e^{-(1+r)^\alpha}. \tag{5}$$

Once the appropriate trial function is selected, the remaining analysis is straightforward in principle—one simply plugs the trial function into $E[\psi]$ and finds an $\alpha>0$ where it is negative. The easiest way to do this is to take the $\alpha\to 0$ limit of the numerator: if this limit is negative, then there must also exist a small $\alpha>0$ where it is negative. This process, which requires some care in taking the limits (limits and integration cannot be interchanged in general), is carried out as follows.

Note that γ is already in polar coordinates and that ψ is normalizable for all such α since ψ_0 is bounded. This trial function has the key physically motivated property that the limit of no localization, i.e., $\alpha\to 0$, gives the unperturbed ground state ψ_0. We write down the first two derivatives of γ for future reference

$$\gamma' \equiv \frac{\partial\gamma}{\partial r} = \hat{r}\cdot\nabla\gamma = -\alpha(1+r)^{\alpha-1}\gamma, \tag{6}$$

$$\gamma'' \equiv \frac{\partial^2\gamma}{\partial r^2} = \alpha(1+r)^{\alpha-2}[\alpha(1+r)^\alpha - \alpha + 1]\gamma. \tag{7}$$

When $H-E_0$ acts only on ψ_0 in $E[\psi]$, the result is zero. The remaining terms in the Rayleigh-quotient numerator, denoted by $U[\psi]$, come from V and derivatives of γ. After some algebraic simplifications, $U[\psi]$ is given by (see the Appendix)

$$\begin{aligned} U[\psi] &= \langle\psi,(H-E_0)\psi\rangle \\ &= \int V|\psi|^2 + \int |\psi_0|^2\left[\frac{1}{2}\nabla^2(\gamma^2) - \gamma\nabla^2\gamma\right]. \end{aligned} \tag{8}$$

Using Eqs. (6) and (7), we obtain

$$\nabla^2\gamma = \alpha(1+r)^{\alpha-2}\left[\alpha(1+r)^\alpha - \alpha - \frac{1}{r}\right]\gamma, \tag{9}$$

$$\nabla^2(\gamma^2) = 2\alpha(1+r)^{\alpha-2}\left[2\alpha(1+r)^\alpha - \alpha - \frac{1}{r}\right]\gamma^2. \tag{10}$$

Plugging these two formulas into $U[\psi]$ results in the concise form

$$U[\psi] = \int V|\psi|^2 + \int |\psi_0|^2\alpha^2(1+r)^{2\alpha-2}\gamma^2. \tag{11}$$

Note that the denominator of $E[\psi]$ (which is $\|\psi\|^2$) is always positive and so does not affect the sign of $U[\psi]$. We want to show that $U[\psi]$, and thus $E[\psi]$, will be negative for some choice of α. This will be done by showing that the term on the right of Eq. (11) tends to zero as $\alpha\to 0$, while $\int V|\psi|^2$ will be negative in this limit. Because $|\psi_0|^2$ is bounded, we have

$$\begin{aligned} &\int |\psi_0|^2\alpha^2(1+r)^{2\alpha-2}\gamma^2 \\ &\quad\le 2\pi\max\{|\psi_0|^2\}\int_0^\infty \alpha^2(1+r)^{2\alpha-2}\gamma^2 r dr \\ &\quad= 2\pi\max\{|\psi_0|^2\}\int_1^\infty \alpha^2 u^{2\alpha-2}\gamma^2(u-1)du \\ &\quad\le 2\pi\max\{|\psi_0|^2\}\int_1^\infty \alpha^2 u^{2\alpha-1}\gamma^2 du, \end{aligned} \tag{12}$$

where we have made the substitution $u=1+r$ and then bounded the integral again. Hence, it suffices to show that the latter integral tends to zero. We calculate this integral explicitly via integration by parts

$$\int_1^\infty \alpha^2 u^{2\alpha-1}\gamma^2 du = -\frac{\alpha}{2}u^\alpha e^{-2u^\alpha}\Big|_1^\infty + \int_1^\infty \frac{\alpha^2}{2}u^{\alpha-1}e^{-2u^\alpha}du = \frac{3\alpha}{4e^2}. \tag{13}$$

Taking the limit as $\alpha\to 0$ yields zero as claimed above.

The leftover $\int V|\psi|^2$ term can be split into two parts. Let $V=V^+-V^-$, where V^+ and V^- are the positive and negative parts of V. Then we have

$$\int |\psi|^2 V = \int |\psi|^2 V^+ - \int |\psi|^2 V^-, \tag{14}$$

where each $|\psi|^2V^\pm$ is a monotonically increasing function as α decreases. This allows us to use Lebesgue's monotone convergence theorem[36] to interchange the limit with the integration, arriving at

$$\lim_{\alpha\to 0}\int |\psi|^2V^\pm = \int \lim_{\alpha\to 0}|\psi|^2V^\pm = e^{-2}\int |\psi_0|^2V^\pm \tag{15}$$

for each part. Putting all the information together, we have

$$\lim_{\alpha\to 0}\left(\int V|\psi|^2 + \int |\psi_0|^2\alpha^2(1+r)^{2\alpha-2}\gamma^2\right) = e^{-2}\int |\psi_0|^2 V < 0 \tag{16}$$

by our main assumption (2). Hence, the variational principle says that there exists an eigenvalue below E_0 for the system and so the theorem is proved.

C. Proof of 1D localization

The case for 1D localization can be proved in an analogous way, with simpler calculations, by using the trial function $\psi_0 e^{-\alpha x^2}$. A closely related result in 1D was presented by Prodan:[4] for any defect potential V satisfying Eq. (2) and for any energy $E<E_0$, Prodan showed that there was some scaling λV, with $\lambda>0$, such that a bound state with energy E exists. Furthermore, the limit $E\to E_0$ was shown to correspond to $\lambda\to 0$, so that an arbitrarily weak potential satisfying (2) must localize a state.

III. 2D LOCALIZATION IN *d* DIMENSIONS

We would like to extend these results to an unperturbed Hamiltonian $H_0=-\nabla^2+V_0$ that is periodic in d dimensions and where the defect potential V is localized along one or two dimensions and is periodic in the other dimensions. A physical example of 2D localization would be a linear defect or "quantum wire" in a 3D crystal, localizing a wave function to propagate along the line, whereas an example of 1D localization in 3D would be a planar defect.

A. Periodicity of V_0 and V

In this case, V_0 is a periodic potential in d dimensions, while V is periodic only in $d-2$ dimensions but localized in the other two. It is convenient to separate the periodic and nonperiodic coordinates of V by writing V as $V(r,\theta,\mathbf{z})$, where $\mathbf{z}\in\mathbb{R}^{d-2}$, which is periodic in $\mathbf{z}$ with some lattice vectors. The first two coordinates are in polar form for convenience in defining what we mean by "localization," which occurs in r only. V_0, on the other hand, is periodic in *all* dimensions, but for convenience, we also write it as $V_0(r,\theta,\mathbf{z})$. A schematic 2D example of such a V and V_0, where V is localized in only one dimension and periodic in one other, is shown in Fig. 2. Note that it is irrelevant in our proof whether V_0 is periodic in any (r,θ) plane, which only occurs if (r,θ) corresponds to a lattice plane[1] of V_0 (such an orthogonal supercell exists only under certain conditions on the lattice vectors[37]).

The corresponding Laplacian for these coordinates is

$$\nabla^2 = \nabla^2_{r,\theta} + \nabla^2_{\mathbf{z}} = \left(\frac{\partial^2}{\partial r^2} + \frac{1}{r}\frac{\partial}{\partial r} + \frac{1}{r^2}\frac{\partial^2}{\partial\theta^2}\right) + \sum_{j=3}^{d}\frac{\partial^2}{\partial x_j^2}. \tag{17}$$

In this section and the next, the symbol $\int f$ will be interpreted as

$$\int f \equiv \int_0^\infty r dr \int_0^{2\pi} d\theta \int_\Omega d^{d-2}\mathbf{z} f(r,\theta,\mathbf{z}), \tag{18}$$

where Ω is the primitive cell in $\mathbf{z}$.

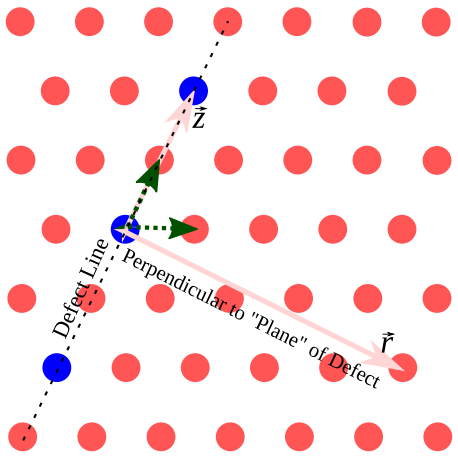

FIG. 2. (Color online) Linear defect for a periodic lattice (red circles) in two dimensions is created by adding a defect potential (blue circles) that is periodic along the $\mathbf{z}$ direction, commensurate with the periodicity of the underlying lattice. This will localize a state with respect to the perpendicular directions, denoted $\mathbf{r}$.

B. Generalized localization below ground state

Due to the periodicity of the potentials, we apply Bloch's theorem to reduce this problem to a unit cell in $\mathbf{z}$, in which case a "localized" mode is a point eigenvalue as in the previous section, and we can prove localization by an identical variational approach.

In particular, all eigensolutions can be chosen in the Bloch form $\psi_0=\psi_0^k(r,\theta,\mathbf{z})e^{i\mathbf{k}\cdot\mathbf{z}}$, where ψ_0^k is periodic in $\mathbf{z}$ with the same periodicity as V (and V_0) and $\mathbf{k}$ is some Bloch wavevector.[1] Substituting this into $H_0\psi=E\psi$, one obtains the "reduced" Hamiltonian

$$H_0^k = -\nabla^2 - 2i\mathbf{k}\cdot\nabla_{\mathbf{z}} + \mathbf{k}^2 + V_0, \tag{19}$$

whose eigenvalues $E(\mathbf{k})$ are the $\mathbf{k}$-dependent band structures. The domain is now only the primitive cell in $\mathbf{z}$ with periodic boundary conditions. For each $\mathbf{k}$, there is a minimum energy, $E_0(\mathbf{k})$, associated with each reduced Hamiltonian H_0^k, i.e., $H_0^k\psi_0^k=E_0(\mathbf{k})\psi_0^k$ for some Bloch wave function ψ_0^k. ($\psi_0^k e^{i\mathbf{k}\cdot\mathbf{z}}$ is merely one of the Bloch wave functions for the underlying potential V_0 such that the projection of the original d-dimensional Bloch wavevector onto the $\mathbf{z}$ dimensions gives $\mathbf{k}$. Therefore, $|\psi_0^k|$ is bounded.) The claim we wish to prove is that the above conditions (just as in Sec. II) guarantee the existence of a bound state with energy below $E_0(\mathbf{k})$ for each $\mathbf{k}$. Moreover, the previous theorem of Sec. II becomes merely the special case when $\mathbf{k}=0$ and there are no extra dimensions of periodicity.

C. Proof of 2D localization in *d* dimensions

The proof follows the same outline as in Sec. II (just replace ψ_0 with ψ_0^k). There are only two small differences. H_0^k in Eq. (19) has the additional terms $2i\mathbf{k}\cdot\nabla_{\mathbf{z}}$ and $\mathbf{k}^2$, but $\nabla_{\mathbf{z}}\gamma=0$ so this additional term only acts on ψ_0^k and is part of the $(H_0-E_0)\psi_0^k=0$ cancellation as before in Eq. (8). The second difference is that these integrals are over $\mathbb{R}^2\times\mathbb{R}^{d-2}$ (where the part over $\mathbb{R}^{d-2}$ is actually just over a primitive

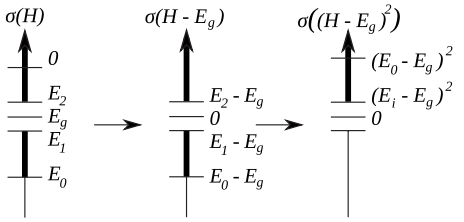

FIG. 3. The leftmost plot is a schematic picture of the spectrum $\sigma(H)$ (with some arbitrary origin 0). The first step is to shift the center of the gap to the zero position by subtracting E_g. The middle plot shows $\sigma(H-E_g)$. Then, the operator is squared and all energies become positive. The final plot shows $\sigma[(H-E_g)^2]$: any eigenvalue introduced into the gap of the original operator will now be an extremal eigenvalue of the shifted and squared operator.

cell), instead of over $\mathbb{R}^2$. This simply means that, in Eq. (12), instead of factoring out $\max\{|\psi_0^k|^2\}$ in the bound, we factor out $\max\{|\psi_0^k|^2\}$ multiplied by the volume of the primitive cell (since the remaining integrand is independent of $\mathbf{z}$).

D. Proof of 1D localization

The analysis is the same except we now have $V_0 = V_0(x,\mathbf{z})$ and $\psi = \psi_0^k e^{i\mathbf{k}\cdot\mathbf{z}} e^{-\alpha x^2}$.

IV. BOUND STATES WITHIN A BAND GAP

Now, we suppose that there is a gap in the spectrum of H_0^k (at a particular $\mathbf{k}$ in d dimensions) from E_1 up to E_2 (with corresponding band-edge Bloch wave functions ψ_1^k and ψ_2^k) as shown in Fig. 1 and we prove sufficient conditions for V to localize a state along one or two dimensions with an energy in the gap (with periodicity as defined in the previous section). Intuitively, if V is mostly negative, then a state will be pulled *down* into the gap from the higher-energy edge E_2. If it is mostly positive instead, then a state will be pushed *up* into the gap from the lower energy edge E_1. On the other hand, if the potential is too strong, then the band-edge state will be pulled or pushed all the way across the gap and will cease to be localized, hinting that we may need some upper bound on the strength of the defect potential even if there is no lower bound. This will be made quantitative in Sec. IV A.

The idea behind the actual proof is to translate it to a simpler problem so that the methods of Sec. II can be used. That is, we must somehow alter the operator H so that we are localizing below the "ground state" of an altered H_0 and can use the variational principle. This is achieved by considering $(H-E_g)^2$, where $E_g=(E_2-E_1)/2$, instead of just H, which transforms the spectrum as shown in Fig. 3. This idea comes from the localization proofs of Ref. 24 as well as from numerical techniques.[26,27] The trial function is motivated once again by Ref. 13.

A. Statement of the theorem

For ease of notation, we will omit the k superscripts in this section. We will prove gap localization under the following conditions. The intuition that V should be mostly negative or positive corresponds to our condition

$$\int V|\psi_1|^2 < 0 \quad \text{and/or} \quad \int V|\psi_2|^2 > 0. \tag{20}$$

The intuition that V cannot be too big corresponds to

$$2|E_i - E_g|\left|\int V|\psi_i|^2\right| > \int V^2|\psi_i|^2, \tag{21}$$

where i depends on which case of Eq. (20) was satisfied; that is, from which band edge we are pulling a localized state. [Conditions (20) and (21) can be merged into a single condition, below, that Eq. (28) is negative.] We also require that V be sufficiently localized in r, corresponding to the condition

$$\int |V||\psi_i|^2 < \infty \quad \text{and} \quad \int V^2|\psi_i|^2 < \infty. \tag{22}$$

These conditions are sufficient to guarantee a bound state *in the gap* for the perturbed Hamiltonian $H=H_0+V$, i.e., a localized state in the (r,θ) dimensions.

B. Proof of 2D localization

Considering one band edge at a time, we will prove localization of a bound state "pulled" from the band edge E_i for $i=1,2$. The proof will be split up into several steps to make it easier to follow. The method employed is the same as that in Sec. II.

The variational principle in this problem can be used after shifting the center of the gap to the zero position and squaring the resulting operator thus making it sufficient to find a normalizable trial function ψ such that

$$\frac{\langle\psi,(H-E_g)^2\psi\rangle}{\langle\psi,\psi\rangle} < (E_i-E_g)^2 \equiv (\Delta E)^2 \tag{23}$$

or, equivalently,

$$\|(H-E_g)\psi\|^2 - (\Delta E)^2\|\psi\|^2 < 0. \tag{24}$$

Consider the trial function $\psi=\gamma\psi_i$, where $\gamma=\exp\{-(1+r)^\alpha\}$ as before. Similar to Sec. II, we will show that this ψ will satisfy condition (24) for some small α providing conditions (20)–(22) are also satisfied.

After some algebraic manipulations and using the fact that $H_0\psi_i=E_i\psi_i$, the left-hand side of Eq. (24) becomes (as shown in more detail in the Appendix)

$$\|(H-E_g)\psi\|^2 - (\Delta E)^2\|\psi\|^2$$
$$= (\Delta E)\int |\psi_i|^2[\nabla^2(\gamma^2) - 2\gamma\nabla^2\gamma], \tag{25}$$

$$+ 2(\Delta E)\int V|\psi|^2 + \|V\psi\|^2$$
$$- \int V[\nabla|\psi_i|^2\cdot\nabla(\gamma^2) - 2|\psi_i|^2\gamma\nabla^2\gamma], \tag{26}$$

$$+ 4\|\nabla\psi_i\cdot\nabla\gamma\|^2 + \|\psi_i\nabla^2\gamma\|^2$$
$$+ 2\int \nabla^2\gamma\nabla|\psi_i|^2\cdot\nabla\gamma. \tag{27}$$

It is a bit cumbersome, but we can separate the integrals into four main groups to avoid calculating each one individually. The first group consists of Eq. (25), but we have already shown that this integral tends to zero as $\alpha \to 0$ in Sec. II [see Eqs. (12) and (13)].

The second group consists of all the integrals which involve V. These are all contained in Eq. (26). Taking the limit as $\alpha \to 0$ is allowed to pass through the integrals because everything is bounded by either a constant times V or a constant times V^2 and our assumption (22) allows us to use Lebesgue's dominated convergence theorem[36] to commute the limits with integration. The limit of these integrals as $\alpha \to 0$ is

$$e^{-2}\int V^2|\psi_i|^2 + 2(E_i - E_g)e^{-2}\int V|\psi_i|^2, \tag{28}$$

where the rightmost integrand of Eq. (26) vanishes for $\alpha \to 0$ because differentiating γ results in at least one α factor. Equation (28) is strictly negative under conditions (20) and (21). Namely, the V^2 term is smaller than the V term by Eq. (21), and Eq. (20) implies that the $(E_i - E_g)V|\psi_i|^2$ integral is negative for either $i=1$ or $i=2$.

We now move on to the final terms: everything in Eq. (27). We wish to show that they all tend to zero as $\alpha \to 0$ and we can do this easily by concentrating on the term that decays *most slowly* with respect to r, so that all the other terms which decay faster clearly go to zero as well (provided that the terms have the same or higher order for the α factor in front). The three integrals of Eq. (27), dropping bounded terms such as $|\psi_i|^2$, are [from left to right in Eq. (27)]:

$$(\nabla\gamma)^2 = \alpha^2(1+r)^{2\alpha-2}\gamma^2, \tag{29}$$

$$(\nabla^2\gamma)^2 = \alpha^2(1+r)^{2\alpha-4}\left[\alpha(1+r)^{\alpha} - \alpha - \frac{1}{r}\right]^2\gamma^2, \tag{30}$$

$$\nabla^2\gamma\nabla\gamma = \alpha^2(1+r)^{2\alpha-3}\left[\alpha(1+r)^{\alpha} - \alpha - \frac{1}{r}\right]\gamma^2. \tag{31}$$

Each has at least an α^2 factor in front. Upon inspection, we find that the most slowly decaying term out of these three is the $(\nabla\gamma)^2$ term, which goes as $1/r^2$ in the limit of $\alpha \to 0$ and its limit is

$$\lim_{\alpha\to 0}\int_0^{\infty} \alpha^2(1+r)^{2\alpha-2}\gamma^2 r dr \to 0, \tag{32}$$

as was already shown in Eqs. (12) and (13). Since Eq. (29) dominates Eq. (27), the other terms are all asymptotically bounded by some constant times this integral and hence must also have integral zero in the limit as $\alpha \to 0$.

What has been shown is that every term from the left-hand side of Eq. (24) vanishes except for the one term (28), which is negative. This establishes the existence of a bound state.

C. Proof of 1D localization

The case for 1D localization can be proved in an analogous way, with simpler calculations, by using the trial function $\psi_0 e^{-\alpha x^2}$ just as before. A closely related result in 1D was presented by Prodan:[4] for any defect potential V satisfying Eq. (20) and for any energy E in the gap, Prodan showed that there was some scaling λV with $\lambda > 0$ such that a bound state with energy E exists. Furthermore, the limit as E approaches the band edge corresponding to Eq. (20) was shown to correspond to $\lambda \to 0$, so that the limit of an arbitrarily weak potential satisfying Eq. (20) must localize a state in the gap.

V. SOME FURTHER GENERALIZATIONS

A. Necessity of periodicity?

In all our derivations, we did not actually explicitly use the fact that V_0 was periodic in the dimensions where localization would take place. All we used were a few of the properties of periodic potentials. These are:

(i) the energies are bounded from below;

(ii) there may be a finite gap inside the continuous spectrum;

(iii) the generalized eigenfunctions corresponding to the infimum or gap-edge energies are bounded (and their derivatives are bounded).

For gap localization, we also assumed that the squared operator was well-defined, making application of this theorem to delta-function potentials (Kronig-Penney models) appear problematic, although it is possible that the difficulty is surmountable with a sufficiently careful limiting process. (In physical contexts, the difference between a theoretical infinite-depth V_0 and a finite-depth approximation seems scarcely relevant.) Also, we assumed coinciding essential spectra for H_0 and H in order to utilize the variational principle. This means that there are some restrictions on how large of a perturbation V can be. However, Weyl's theorem states one sufficient condition,[31,32] which appears to be satisfied for most physically interesting V's; see also the comments after Eq. (3).

Assumption (iii) may be more challenging to prove for nonperiodic potentials. We assumed that the energies at the infimum of the spectrum and/or the edges of spectral gaps correspond to eigenvalues with bounded generalized eigenfunctions (ψ_0 or ψ_i). (This corresponds to the requirement of a regular ground state in Ref. 15.) For periodic potentials, the existence of a band-edge solution for V_0 follows from the well-known continuity of the band diagram as a function of the Bloch wavevector $\mathbf{k}$. For nonperiodic potentials V_0, however, this is not necessarily true. For example, for the 1D half-well potential $V_0(x)=\infty$ for $x<0$ and $=0$ for $x>0$, the eigenfunctions $\sin(\kappa x)$ do not have a nonzero infimum-energy solution for $\kappa \to 0$ and correspondingly it is well known that any perturbation V must exceed some threshold depth before a bound state appears in that case—this is mathematically equivalent to the appearance of an odd bound-state solution $\psi(-x)=-\psi(x)$ for $V_0=0$ and an even perturbation $V(-x)=V(x)$, which requires a threshold depth since the lowest-energy bound state in that case is even. It is not clear to us under what conditions the requisite infimum or gap-edge solutions exist for more general potentials, such as quasiperiodic potentials V_0, although some examples are given in Ref. 15.

B. Degeneracy at the band edges

As mentioned earlier, it could happen that there are multiple (degenerate) linearly independent ψ_i's corresponding to a given energy E_i at an edge of the gap and/or at the infimum of the spectrum. Our proof in the preceding sections is unaffected—there must still be *at least one* bound state localized by a suitable V, as long as the requisite conditions [Eq. (2) or (20)–(22)] hold for at least one of the degenerate ψ_i wave functions. Intuitively, however, one might expect to be able to prove a stronger theorem in this case—if E_i is D-fold degenerate, can one show that D-localized states are formed by a suitable V?

To prove the existence of more than one localized state, we can employ a generalization of the min-max theorem. For a single localized state, our proof involved the fact that the ground-state eigenvalue of a Hermitian operator O is bounded above by the Rayleigh quotient $Q[\psi] = \langle \psi, O\psi \rangle / \langle \psi, \psi \rangle$ for any ψ. The generalization of this fact is that the nth eigenvalue λ_n is bounded above by[19]

$$\lambda_n \leq \sup_{\psi \in S_n} Q[\psi], \tag{33}$$

where S_n is any n-dimensional subspace of the Sobolev space for O. We then wish to show that $\lambda_n < b$ for some bound b: $O = H - E_0$ and $b = 0$ for localization below the infimum of the spectrum or $O = (H - E_g)^2$ and $b = (E_i - E_g)^2$ for localization in the gap from edge i. This is equivalent to proving that the Hermitian form $B[\psi, \phi] = \langle \psi, (O-b)\phi \rangle$ is negative definite for *some* n-dimensional subspace S_n (i.e., $B[\psi, \psi] < 0$ for all $\psi \in S_n$).

If E_i is D-fold degenerate, with degenerate generalized eigenfunctions ψ_i^ℓ for $\ell = 1, \ldots, D$, then the analog of our previous approach is to form the trial functions $\psi^\ell = \gamma \psi_i^\ell$ (whose span is a subspace S_D), compute the $D \times D$ matrix $\mathcal{B}_{\ell\ell'} = B[\psi^\ell, \psi^{\ell'}]$, and check whether it is negative definite as $\alpha \to 0$. We wish to find the largest subspace S_n of S_D for which B is negative definite, which corresponds to the number n of negative eigenvalues of $\mathcal{B}$: this will be the number n of localized states that are guaranteed by the theorem.

For localization below the infimum of the spectrum by a V satisfying Eq. (3), following exactly the same steps as in Sec. III, proving that $B[\psi, \phi]$ is negative definite in this subspace reduces to a generalization of condition (2). Specifically, showing $B[\psi, \psi] < 0$ in the subspace for $\alpha \to 0$ reduces, via Eq. (16), to showing that $\int V|\psi|^2 < 0$ for every ψ in the subspace. In other words, the Hermitian form $A[\psi, \phi] = \langle \psi, V\phi \rangle$ must be negative definite in S_n. In the $\alpha \to 0$ limit, this corresponds to checking the eigenvalues of the $D \times D$ matrix $\mathcal{A}_{\ell\ell'} = A[\psi_0^\ell, \psi_0^{\ell'}]$: the number of negative eigenvalues of $\mathcal{A}$ is precisely the dimension of the largest negative-definite subspace S_n and hence is the number of bound states that are guaranteed to be localized below the ground state of V_0. If we happen to have a strictly nonpositive $V \leq 0$, then the Hermitian form A is automatically negative definite and we are guaranteed D localized modes.

For localization in a gap by a V satisfying Eq. (22), pulling states from band edge i, following exactly the same steps as in Sec. IV, one finds that $B[\psi, \phi]$ being negative definite reduces to a generalization of condition (28): the Hermitian form $G[\psi, \phi] = \langle \psi, [V^2 + 2(E_i - E_g)V]\phi \rangle$ must be negative definite in S_n. As above, this simplifies for $\alpha \to 0$ to counting the number of negative eigenvalues of the $D \times D$ matrix $\mathcal{G}_{\ell\ell'} = G[\psi_i^\ell, \psi_i^{\ell'}]$. The number of negative eigenvalues is the number of solutions that are guaranteed to be localized from band edge i. If V has sign everywhere opposite to $E_i - E_g$ and is sufficiently small (to overwhelm the V^2 term), then D eigenstates will be localized from this band edge.

This analysis appears to be closely related to the asymptotic technique of Ref. 28, which also relates a number of bound modes to the number of eigenvalues of a given sign of a small matrix, via the Birman-Schwinger principle in the limit of weak perturbations, but that work only explicitly considered localization below the ground state of translation-invariant elliptic and Schrödinger-type unperturbed operators.

VI. CONCLUDING REMARKS

Although the existence of localized solutions from defects in periodic potentials and the effective-mass analogy with the vacuum case are well known as a practical computational and experimental matter, it is gratifying to have a general, explicit proof that localization in one and two dimensions occurs in a similar manner to localization in vacuum. A number of directions suggest themselves for future research. Although the simplicity of an elementary proof based on the min-max or variational theorem has its own appeal, the application of more sophisticated methods such as those of Ref. 11 may reveal additional information about the nature of the localized state (such as its asymptotic localization length) that cannot be gleaned from a simple variational analysis. We would also like to transfer these results from the Schrödinger (scalar) picture to the Maxwell (vector) one, in the context of localization in band gaps of photonic crystals such as photonic-crystal fibers.[2] A similar generalization to electromagnetism was already obtained for localization below the infimum of the spectrum (corresponding to total internal reflection in Maxwell's equations) for nonperiodic[19] and periodic[20] media and in photonic band gaps for sufficiently large defects.[24,38]

For localization in gaps, we should remark that the condition (21) on the size of the perturbation V is somewhat unsatisfying. Intuitively, a "small" perturbation V could be one where either $|V|$ is small at every point or where $|V|$ is not small but the support of V is small. The latter case, however, of a large $|V|$ with a small integral violates our smallness condition (21) for a sufficiently large $|V|$ no matter how small the support might be. This does not mean that there are no localized states in that limit—our proof only gives a sufficient condition for localization, not a necessary condition—but it suggests that some reformulation to handle this physically interesting possibility might be desirable.

ACKNOWLEDGMENTS

We are grateful to Fritz Gesztesy, Dirk Hundertmark, Patrick Lee, Leonid Levitov, Emil Prodan, Alexander Push-

nitski, Barry Simon, and Michael Weinstein for helpful discussions and references to the literature. This work was supported in part by the Macaulay Honors College at the City University of New York.

APPENDIX

Here we will provide more details for the calculations of Secs. II and IV. Let ψ be an eigenstate multiplied by a *real* function γ whose derivatives decay quickly enough, i.e., $\psi = \psi_0 \gamma$, where and γ is real (we can think of the example used in this paper) and ψ_0 is bounded and satisfies $H_0\psi_0 = E_0\psi_0$. We will explore some of the terms in both $\langle \psi, (H-E_0)\psi\rangle$ and $\langle (H-E_g)\psi, (H-E_g)\psi\rangle$. Our goals in the above sections were to show that most terms tended to zero. In order for this to be apparent, we had to perform several algebraic manipulations of the integrals that are given here more explicitly.

1. Calculations for Sec. II

We will derive the general formula used in Sec. II first. We have

$$(H-E_0)\psi = (-\nabla^2\psi_0 + V_0\psi_0)\gamma - E_0\psi_0\gamma + V\psi - 2\nabla\psi_0\cdot\nabla\gamma - \psi_0\nabla^2\gamma = V\psi - 2\nabla\psi_0\cdot\nabla\gamma - \psi_0\nabla^2\gamma, \quad \text{(A1)}$$

since $(-\nabla^2 + V_0)\psi_0 = E_0\psi_0$, so that $U[\psi] = \langle \psi, (H-E_0)\psi\rangle$ is

$$U[\psi] = \int V|\psi|^2 - 2\int \gamma\psi_0^*\nabla\psi_0\cdot\nabla\gamma - \int |\psi_0|^2\gamma\nabla^2\gamma. \quad \text{(A2)}$$

Because $U[\psi]$ must always be real and γ is assumed to be real, the imaginary part of $\psi_0^*\nabla\psi_0$ must integrate to zero. Therefore, we can replace $2\psi_0^*\nabla\psi_0$ by $\psi_0^*\nabla\psi_0 + \psi_0\nabla\psi_0^*$ in the integrand. We can also use the identity $\gamma\nabla\gamma = \frac{1}{2}\nabla(\gamma^2)$ so that Eq. (A2) becomes

$$\int V|\psi|^2 - \frac{1}{2}\int \nabla|\psi_0|^2\cdot\nabla(\gamma^2) - \int |\psi_0|^2\gamma\nabla^2\gamma. \quad \text{(A3)}$$

We now rewrite middle term in another manner in order to eliminate the slowly decaying second derivative of γ using integration by parts

$$\int \nabla|\psi_0|^2\cdot\nabla(\gamma^2) = \int_\partial |\psi_0|^2\nabla(\gamma^2) - \int |\psi_0|^2\nabla^2(\gamma^2) \quad \text{(A4)}$$

$$= -\int |\psi_0|^2\nabla^2(\gamma^2), \quad \text{(A5)}$$

where "$\int_\partial$" stands for the boundary integral. The boundary integral is zero because the γ term and it first derivative decay fast enough and ψ_0 is bounded. Therefore, all we have to show in Sec. II is that

$$\int V|\psi|^2 + \frac{1}{2}\int |\psi_0|^2\nabla^2(\gamma^2) - \int |\psi_0|^2\gamma\nabla^2\gamma < 0 \quad \text{(A6)}$$

for some sufficiently small choice of the parameter α.

2. Calculations for Sec. IV

We will now derive the general formula used in Sec. IV. The same assumptions hold as in the previous section, but with minor modifications from Sec. IV, for example, $\psi = \psi_i\gamma$, where $i=1,2$, signifies the lower or upper edge of the band gap, respectively. First, we calculate $(H-E_g)\psi$,

$$(H-E_g)\psi = (\Delta E)\psi + V\psi - 2\nabla\psi_i\cdot\nabla\gamma - \psi_i\nabla^2\gamma, \quad \text{(A7)}$$

where $\Delta E = E_i - E_g$ as in Sec. IV. Then the generalized equation for $\|(H-E_g)\psi\|^2$ is given by (employing some of the trivial simplifications from the previous section)

$$(\Delta E)^2\|\psi\|^2 + 2(\Delta E)\int V|\psi|^2 + \|V\psi\|^2 \quad \text{(A8)}$$

$$- \int V\nabla|\psi_i|^2\cdot\nabla(\gamma^2) - 2\int V|\psi_i|^2\gamma\nabla^2\gamma \quad \text{(A9)}$$

$$- 2(\Delta E)\int |\psi_i|^2\gamma\nabla^2\gamma \quad \text{(A10)}$$

$$+ \|\psi_i\nabla^2\gamma\|^2 + 2\int \nabla^2\gamma\nabla|\psi_i|^2\cdot\nabla\gamma \quad \text{(A11)}$$

$$- (\Delta E)\int \nabla|\psi_i|^2\cdot\nabla(\gamma^2) + 4\|\nabla\psi_i\cdot\nabla\gamma\|^2. \quad \text{(A12)}$$

Notice that the first term in Eq. (A12) is similar to the middle term in Eq. (A3) from the previous section, so the same analysis shows that the terms of Eq. (A12) become

$$+ (\Delta E)\int |\psi_i|^2\nabla^2(\gamma^2) + 4\|\nabla\psi_i\cdot\nabla\gamma\|^2. \quad \text{(A13)}$$

In Eq. (25), this first term of Eq. (A13) is combined with term (A10). Several of the other terms are rearranged to make the presentation more concise.

*stevenj@math.mit.edu

[1] N. W. Ashcroft and N. D. Mermin, *Solid State Physics* (Harcourt, Orlando, FL, 1976).

[2] J. D. Joannopoulos, S. G. Johnson, J. N. Winn, and R. D. Meade, *Photonic Crystals: Molding the Flow of Light*, 2nd ed. (Princeton University Press, Princeton, NJ, 2008).

[3] N. F. Mott and R. W. Gurney, *Electronic Processes in Ionic Crystals* (Dover, New York, 1964).

[4] E. Prodan, Phys. Rev. B **73**, 085108 (2006).

[5] P. A. Deift and R. Hempel, Commun. Math. Phys. **103**, 461

(1986).
[6] F. Gesztesy and B. Simon, Commun. Math. Phys. **116**, 503 (1988).
[7] R. Hempel, J. Math. Anal. Appl. **169**, 229 (1992).
[8] O. L. Safronov, J. Math. Anal. Appl. **260**, 641 (2001).
[9] D. Hundertmark and B. Simon, J. Math. Anal. Appl. **340**, 892 (2008).
[10] O. Post, Math. Nachr. **261-262**, 141 (2003).
[11] B. Simon, Ann. Phys. **97**, 279 (1976).
[12] E. N. Economou, *Green's Functions in Quantum Physics*, 3rd ed. (Springer, Heidelberg, 2006).
[13] K. Yang and M. de Llano, Am. J. Phys. **57**, 85 (1989).
[14] D. Hundertmark, Proc. Symp. Pure Math. **76**, 463 (2007).
[15] R. L. Frank, B. Simon, and T. Weidl, Commun. Math. Phys. **282**, 199 (2008).
[16] L. D. Landau and E. M. Lifshitz, *Quantum Mechanics: Non-Relativistic Theory* (Pergamon Press, Oxford, 1977).
[17] S. Brandt and H. D. Dahmen, *The Picture Book of Quantum Mechanics*, 3rd ed. (Springer, Heidelberg, 2000).
[18] H. Picq, Ph.D. thesis, Université de Nice, 1982.
[19] A. Bamberger and A. S. Bonnet, SIAM J. Math. Anal. **21**, 1487 (1990).
[20] K. K. Y. Lee, Y. Avniel, and S. G. Johnson, Opt. Express **16**, 9261 (2008).
[21] D. Damanik, D. Hundertmark, R. Killip, and B. Simon, Commun. Math. Phys. **238**, 545 (2003).
[22] F. Gesztesy and B. Simon, Trans. Am. Math. Soc. **335**, 329 (1993).
[23] R. Hempel, Contemp. Math. **458**, 393 (2008).
[24] P. Kuchment and B. Ong, in *Waves in Periodic and Random Media*, Contemporary Mathematics Vol. 339 (American Mathematical Society, Providence, 2003), pp. 105–115.
[25] *Templates for the Solution of Algebraic Eigenvalue Problems: A Practical Guide*, edited by Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. Van Der Vorst (SIAM, Philadelphia, 2000).
[26] S. Johnson and J. Joannopoulos, Opt. Express **8**, 173 (2001).
[27] L.-W. Wang and A. Zunger, J. Chem. Phys. **100**, 2394 (1994).
[28] T. Weidl, Commun. Partial Differ. Equ. **24**, 25 (1999).
[29] A. Pushnitski, F. Gesztesy, and B. Simon (private communication).
[30] B. Ilan and M. I. Weinstein, Multiscale Model. Simul. (to be published).
[31] P. D. Hislop and I. M. Sigal, *Introduction to Spectral Theory With Applications to Schrödinger Operators*, Applied Mathematical Sciences Vol. 113 (Springer, New York, 1996).
[32] M. Schechter, *Operator Methods in Quantum Mechanics* (North-Holland, Amsterdam, 1981).
[33] D. J. Griffiths, *Introduction to Quantum Mechanics*, 2nd ed. (Benjamin Cummings, New York, 2004).
[34] R. Shankar, *Principles of Quantum Mechanics*, 2nd ed. (Springer, Heidelberg, 1994).
[35] M. Reed and B. Simon, *Methods of Modern Mathematical Physics: Vol. IV Analysis of Operators* (Academic Press, London, 1978).
[36] W. Rudin, *Real and Complex Analysis*, 3rd ed. (McGraw-Hill, New York, 1987).
[37] R. Bucksch, J. Appl. Crystallogr. **6**, 400 (1973).
[38] D. Miao, J. Math. Phys. **49**, 063511 (2008).